

An Adaptive Spread Spectrum (SS) Synchronous Data Hiding Strategy for Scalable 3D Terrain Visualization

K. Hayat^a, W. Puech^{ab} and G. Gesquière^c

^aLIRMM, UMR CNRS 5506, University of Montpellier II
161, rue Ada, 34392 MONTPELLIER CEDEX 05, FRANCE

^bUniversity of NIMES, Place G. Péri, 30021 NIMES CEDEX 1, FRANCE

^cLSIS, UMR CNRS 6168, Aix-Marseille University
IUT, rue R. Follereau, 13200 ARLES, FRANCE

ABSTRACT

The diversity of clients in today's network environment compels us to think about solutions that more than satisfy their needs according to their resources. For 3D terrain visualization this translates into two main requirements, namely the scalability and synchronous unification of a disparate data that requires at least two files, the texture image and its corresponding digital elevation model (DEM). In this work the scalability is achieved through the multiresolution discrete wavelet transform (DWT) of the JPEG2000 codec. For the unification of data, a simple DWT-domain spread spectrum (SS) strategy is employed in order to synchronously hide the DEM in the corresponding texture while conserving the JPEG2000 standard file format. Highest possible quality texture is renderable due to the reversible nature of the SS data hiding. As far as DEM quality is concerned, it is ensured through the adaptation of synchronization in embedding that would exclude some highest frequency subbands. To estimate the maximum tolerable error in the DEM according to a given viewpoint, a human visual system (HVS) based psycho-visual analysis is being presented. This analysis is helpful in determining the degree of adaptation in synchronization.

Keywords: Spread Spectrum, Scalable Data Hiding, Adaptive Synchronization, DEM, Texture, 3D Terrain Visualization, JPEG2000, DWT, Approximation Image, Multi-resolution, GIS.

1. INTRODUCTION

Terrain visualization in 3D involves mapping of an aerial photograph, called texture image, onto a regular triangulated surface created from a corresponding set of coefficients, called DEM.^{1,2} Each of the DEM coefficient represents the altitude of a particular square block of texture pixels. The quality of visualization depends on the precision and accuracy of both the texture and its DEM and usually involves a lot of data even for a small district. For interactive visualization, handling of such a huge data requires a lot of memory, bandwidth and processing. In a contemporary client/server environment one comes across a variety of clients in terms of network connection, memory and computing resources, access rights, preferences, distance from the server and many others. The wide spectrum of clients compels us to devise some automated scalable data compression mechanism that is ideally scalable to each hierarchy of clients. The JPEG2000³ standard serves this purpose since it is DWT-based and the multi-resolution property of the DWT can give us the required scalability. In doing so, the JPEG2000 file format should better be conserved instead of proposing some new proprietary format. Hence in place of using two disparate DEM and texture files it is advisable to interleave them into a single file which can then be converted to JPEG2000 format. For this purpose one can take a cue from the technique of data hiding and embed the smaller DEM data inside the larger texture image. The embedding scheme must have to address the issue of correspondence between the texture and its DEM. In other words the integration of texture and its DEM into a single file must be synchronous.

Further author information:

K.H.: E-mail: khizar.hayat@lirmm.fr, Telephone: +33 (0)4 67 41 85 64

W.P.: E-mail: william.puech@lirmm.fr, Telephone: +33 (0)4 67 41 86 85

G.G.: E-mail: gilles.gesquiere@lsis.org, Telephone: +33 (0)4 90 52 24 10

In our previous efforts, LSB-based data hiding strategy was employed to synchronously unify lossless wavelet transformed DEM in the Y/Cr/Cb plane of the corresponding texture in the lossless,⁴ or lossy⁵ DWT domain, usually from the JPEG2000 pipeline. We improve upon these in our present work by specially focusing on the quality of both the texture and the DEM. To ameliorate the final DEM quality, one feels it prudent to adapt the synchronization by harmonizing the embedding with a subset of luminance/chrominance subbands of texture, rather than the whole, as was done in Ref. 6. Because of the sensitive nature of DEM it will be expedient to explore some way to avoid too much loss in its quality during the course of the method. The significant difference in the sizes of DEM and texture allows us the luxury to use much smaller texture blocks for embedding. At present, block size of the texture is much higher than the size of a single DEM coefficient in bits. Since the number of required blocks is fixed, size reduction will automatically exclude some higher subbands from the embedding process. The strategy has to be then to use a subset rather than all the texture subbands of wavelet decomposition for embedding. The level of DWT decomposition of the DEM would then obviously be lower than that of the texture. Since the main information of the image is concentrated in lower subbands than the higher ones, any noise in the former would affect the image quality more. Since the packet order in a JPEG2000 stream is from low to high subbands, one is compelled to prefer the lower subbands for embedding. Sight must not be lost of the fact that lower subbands are smaller in size than the higher ones. we opt for a blind strategy that may not necessarily be non-removable. In our approach we propose to employ a spread spectrum (SS) data hiding strategy pioneered by Cox *et al.*⁷ The SS methods offer high robustness at the expense of cover quality but this quality loss is reversible since the embedded data can be removed after recovery.

The adaptive synchronization proposed in this paper is helpful in improving the quality of the DEM approximation for a given texture approximation. The DEM error can thus be reduced at the expense of texture quality but since the data hiding step is reversible, high texture quality is still realizable. To ascertain the maximum tolerable DEM error, from a given viewpoint, a psycho-visual analysis is given that is supported by a mathematical treatment.

The rest of the paper is arranged as follows. Section 2 briefly presents the previous work relevant to our approach. Our method is explained in Section 3 while the results which we obtained are elaborated in Section 4. Section 5 concludes this paper.

2. STATE OF THE ART

Many methods have been proposed in the literature for wavelet-based data hiding but few of these are compatible with the JPEG2000 scheme. Data hiding methods for JPEG2000 images must process the code blocks independently⁸ and that is why methods like inter-subband embedding⁹ and those based on hierarchical multi-resolution relationship¹⁰ are not recommended. There are methods^{11,12} for embedding invisible watermarks by adding pseudo-random codes to large coefficients of the high and middle frequency bands of DWT but the methods have the disadvantage of being non-blind. The blind scheme proposed by Su *et al.*¹³ is to integrate data hiding with the Embedded Block Coding with Optimized Truncation (EBCOT) and embed data during the formation of compressed bit stream. One particular technique¹⁴ embeds watermark in the JPEG2000 pipeline after the stages of quantization and region of interest (ROI) scaling but before the entropy coding. A window sliding approach is adopted for embedding and for the sake of reliability the finest resolution subbands are avoided while the lowest frequencies carry higher payload.

Piva *et al.* have proposed an authentication scheme that embeds an image digest in a subset of the subbands from the DWT domain.¹⁵ The image digest is derived from the DCT of the level 1 DWT *LL* subband of the image. The resultant DCT coefficients are scaled down by quantization and ordered from most to least significant through a zig-zag scan. A most significant subset, after discarding the DC coefficient, is quadruplicated for redundancy and then rescaled and scrambled by using two different keys. This gives the message which is substituted to the subbands selected from a set obtained by the further wavelet decomposition of the level 1 *HL* and *LH* subbands of the original image.

One blind method¹⁶ transforms the original image by one-level wavelet transform and sets the three higher subbands to zero before inverse transforming it to get the modified image. The difference values between the original image and the modified image are used to ascertain the potential embedding locations of which a subset

is selected pseudo-randomly for embedding. The method of Kong *et al.* embeds watermark in the weighted mean of the wavelets blocks, rather than in the individual coefficient, to make it robust and perceptually transparent.¹⁷ While explaining their method of embedding biometric data in fingerprint images, Noore *et al.* argue against the modification of the lowest subband to avoid degradation of the reconstructed image as most of the energy is concentrated in this band.¹⁸ Instead they propose to redundantly embed information in all the higher frequency subbands.

While using the invariance of the histogram shape Xiang and Kim¹⁹ rely on time-frequency localization property of DWT to propose a watermarking scheme that is resistant to geometric deformations. A geometrically invariant watermark is embedded into the low-frequency sub-band of DWT in such a way that the watermark is not only invariant to various geometric transforms, but also robust to common image processing operations.

Based on the significant difference of wavelet coefficient quantization a blind algorithm²⁰ group every seven non-overlap wavelet coefficients of the host image into a block. The largest two coefficients in a given block are referred to as significant coefficients and their difference as significant difference. The local maximum wavelet coefficient in a block is quantized by comparing the significant difference value in a block with the average significant difference value in all blocks. The maximum wavelet coefficients are so quantized that their significant difference between watermark bit 0 and watermark bit 1 exhibits a large energy difference which can be used for watermark extraction. During the extraction, an adaptive threshold value is designed to extract the watermark from the watermarked image under different attacks. To determine the watermark bit the adaptive threshold value is compared to the block-quantized significant difference.

The JPEG2000-based image authentication method of Schauweg²¹ wherein all coefficients of the wavelet decomposition are protected using their extended scalar quantization and hashing scheme. The method is claimed to be robust to JPEG as well as JPEG2000 re-compression as they employ error correction coding that yields impressive robustness improvements of the embedded image content dependent signature without raising a security gap. The process involves feature extraction by wavelets to result in digital signature which, after encryption and error correction coding, is embedded as a removable watermark using the well-known quantization index modulation technique called dither modulation. The embedded watermark information is removable during the decompression process which is important for the improved image quality in the context of visualization.

Agreste *et al.*²² put forward a strong wavelet-based watermarking algorithm, called WM2.0. WM2.0 embeds the watermark into high frequency DWT components of a specific sub-image and it is calculated in correlation with the image features and statistical properties. Watermark detection applies a re-synchronization between the original and watermarked image. The correlation between the watermarked DWT coefficients and the watermark signal is calculated according to the Neyman-Pearson statistic criterion just like the blind chaotic method²³ of DWT oriented watermarking.

Uccheddu *et al.*²⁴ adopt a wavelet framework in their blind watermarking scheme for 3D models under the assumption that the host meshes are semi-regular ones paving the way for a wavelet decomposition and embedding of the watermark at a suitable resolution level. For the sake of robustness the host mesh is normalized by a Principal Component Analysis (PCA) before embedding. Watermark detection is accomplished by computing the correlation between the watermark signal and the "to-be-inspected" mesh.

Yin *et al.*²⁵ adopt Guskov's multiresolution signal processing method for meshes and use his 3D non-uniform relaxation operator to construct a Burt-Adelson pyramid for the mesh, and then watermark information is embedded into a suitable coarser mesh. The algorithm is integrable with the multiresolution mesh processing toolbox and watermark detection requires registration and resampling to bring the attacked mesh model back into its original location, orientation, scale, topology and resolution level.

A good and elaborate application of wavelets to real-time terrain visualization can be found in Ref. 26 which rely on restricted quadtree triangulation for surface approximation. The focus of this work is however the third dimension, i.e. DEM and suffers from the lack of integration with the texture. However, there are already efforts like GeoJP2²⁷ and GMLJP2²⁸ for the integration of geographical data with the related aerial photographs. The GeoJP2 is a GeoTIFF-inspired method for adding geospatial metadata to a JPEG2000 file. The GeoTIFF specification* defines a set of TIFF tags provided to describe all cartographic information associated with TIFF

*<http://www.remotesensing.org/geotiff/spec/contents.html>

imagery that originates from satellite imaging systems, scanned aerial photography, scanned maps and DEM. Its aim is to allow means for tying a raster image to a known model space or map projection, and for describing those projections. GeoTIFF does not intend to become a replacement for existing geographic data interchange standards. GeoTIFF uses a small set of reserved TIFF tags to store a broad range of georeferencing information, catering to geographic as well as projected coordinate systems needs. It uses numerical codes to describe projection types, coordinate systems, datums and ellipsoids.

The additions made to the box-based JPEG2000 format are two UUID boxes, namely, the GeoTIFF box and the optional world file box. The former contains a degenerate GeoTIFF file and the resultant JPEG2000 file have the same level of geospatial metadata as is provided by the GeoTIFF standard. The mechanism is simple using the widely supported GeoTIFF implementations but the introduction of new UUID boxes have the disadvantage that there is an increase in the original JPEG2000 file size. The GMLJP2 envisages the use of the Geography Markup Language (GML) within the XML boxes of the JPEG 2000 data format in the context of geographic imagery. A minimally required GML definition is specified for geo-referencing images while also giving guidelines for encoding of meta-data, features, annotations, styles, coordinate reference systems, and units of measure as well as packaging mechanisms for both single and multiple geographic images. DEMs are treated the same way as other image use cases whereas coordinate reference system definitions are employed using a dictionary file. Thus DEM is either provided as TIFF file and its name is inserted between proper GML tags or its points are directly inserted into the GMLJP2 file. In the former case, there is no reduction in the number of files whereas in the latter case the amount of data is increased.

3. THE PROPOSED SS METHOD

In this section, we present, a new method for a scalable transfer and online visualization of textured 3D terrain data. A brief overview of the method is described in Section 3.1. Section 3.2 depicts the viewpoint analysis for the adaptation of synchronization while the embedding phase is explained in Section 3.3. The decoding and reconstruction procedures are given in Section 3.4 and Section 3.5, respectively.

3.1 Overview of the method

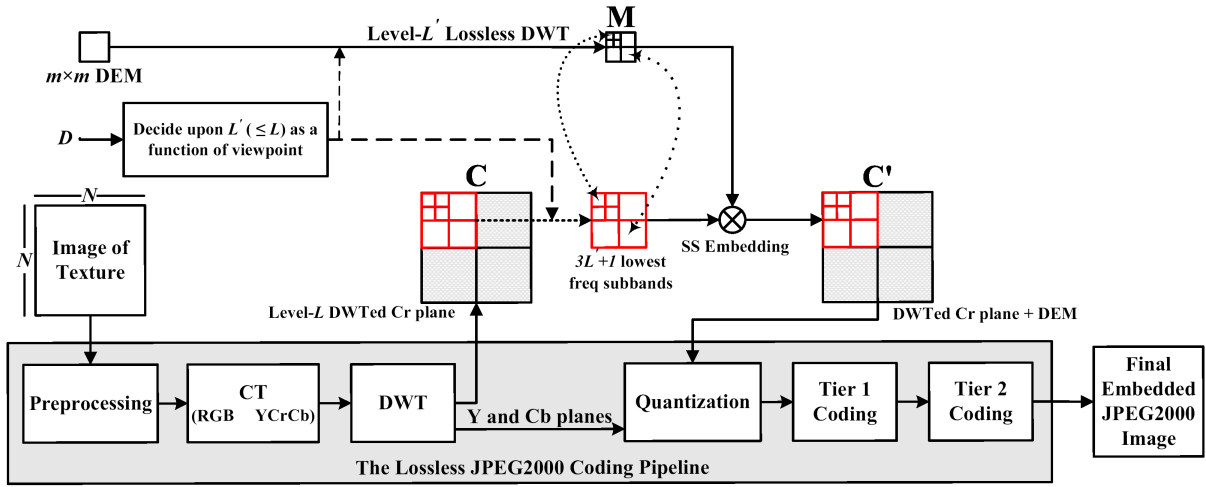


Figure 1: Global overview of the encoding process when using the Cr plane for embedding

Suppose a $N \times N$ texture image has $m \times m$ altitudes in its DEM. In the spatial domain, let one DEM coefficient corresponds to a $t \times t$ pixel block of the related texture, where $t = \frac{N}{m}$. Suppose the texture is to be JPEG2000 coded at DWT decomposition level L , implying $R = L + 1$ resolutions. Let us apply lossless DWT to DEM coefficients at level L' , where $L' \leq L$. For embedding we interrupt the JPEG2000 coding, of the texture, after the DWT step, as illustrated in Fig. 1, and get one of the transformed YCrCb components, say Cr, as a

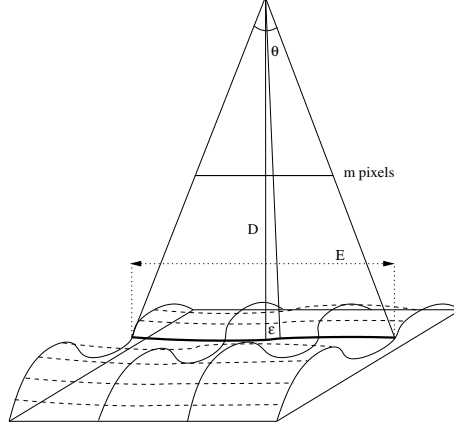


Figure 2: Visualization of a DEM from a viewpoint.

carrier for embedding. The carrier (C) is partitioned into $m \times m$ equal-sized blocks, $B_{i,j}$, with size dependent on the value of L' . If $L' = L$ then whole of the selected component is the carrier (C) and block size remains to be $t \times t$, for embedding, since all the subbands are included for possible data insertion. Otherwise, for $L' < L$, only a subset subbands - the lowest $3L' + 1$ of the original $3L + 1$ after excluding the remaining $3(L - L')$ higher frequency subbands - constitute C and $B_{i,j}$ has a reduced size of $t/2^{(L-L')} \times t/2^{(L-L')}$. Care must be taken of the fact that block size must be large enough to reliably recover the embedded data after correlation.

3.2 Degree of adaptation in synchronization

Our strategy is aimed at a good 3D visualization as a function of the network connection and the computing resources of the client. The proposed adaptive approach to embed the DEM in the DWT texture is a function of the distance D between the viewpoint and the DEM. The DEM quality is evaluated with the root mean square error (RMSE) in meters (m). As pointed out in some recently published works,^{29,30} even today the acquisition of the DEM is error prone and it is difficult to get a RMSE less than 1 meter. To calculate the maximum acceptable RMSE, for an optimal 3D visualization, we rely on the distance D between the viewpoint and the DEM, as illustrated in Fig. 2, and the visual acuity (VA) of the human visual system (HVS). Visual acuity is the spatial resolving capacity of the HVS. It may be thought of as the ability of the eye to see fine details. There are various ways to measure and specify visual acuity, depending on the type of acuity task used. VA is limited by diffraction, aberrations and photoreceptor density in the eye.^{31,32} In this paper for the HVS we assume that the VA corresponds to an arc θ_{VA} of 1 minute ($1' = 1/60^\circ$). Then for a distance D the level of detail (LOD) is:

$$LOD = 2 \times D \times \tan(\theta_{VA}). \quad (1)$$

For example if $D = 3 \text{ m}$ then $LOD = 87 \times 10^{-4} \text{ m}$ and if $D = 4 \text{ km}$ then $LOD = 1.164 \text{ m}$. For our application, illustrated in Fig. 2, if we want to see all the DEM we have a relation between D and the size of the DEM (width = E in meters (m)):

$$E = 2 \times D \times \tan(\theta/2). \quad (2)$$

Usually the field of view of the HVS is $\theta = 60^\circ$. For example if $E = 3200 \text{ m}$ then $D = 2771.28 \text{ m}$.

For the 3D visualization, we should take into account of the resolution of the screen in pixels. As illustrated in Fig. 2, if we have an image or a screen with m pixels for each row, then the LOD ϵ in m is:

$$\epsilon = \frac{E/2}{\tan(\theta/2)} \times \tan(\theta/m). \quad (3)$$

With $\theta = 60^\circ$, $E = 3200\text{ m}$ and a resolution of 320 pixels (for a PDA for example), we have a LOD $\epsilon = 9.07\text{ m}$. Then, in this context, for our application we can assume that a RMSE near 9 m is acceptable for the DEM. Notice that we generate a conservative bound by placing an error of the maximum size as close to the viewer as possible with an orthogonal projection of the viewpoint on the DEM. We also assume that the DEM is globally flat and that the accuracy between the center and the border of the DEM is the same. In reality, the analysis should be different and part of the border should be cut as explained in Ref. 33 as a particular case of a cylinder. Anyhow, the value of D would help us to reach a decision about the value of L' .

3.3 The embedding step

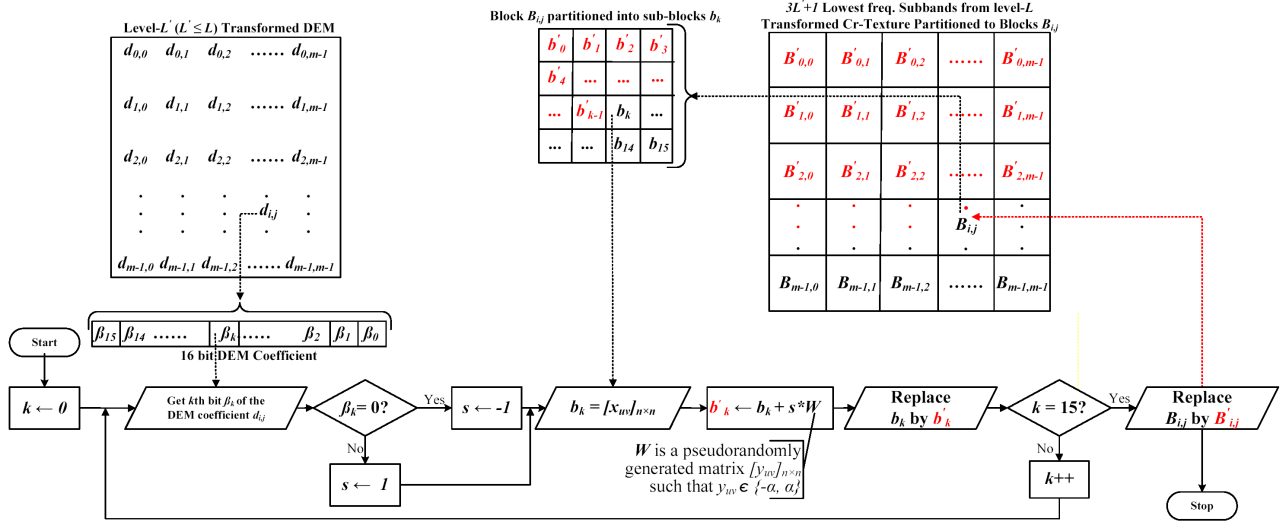


Figure 3: Flowchart showing the DWT domain SS embedding of the DEM coefficient $d_{i,j}$ in the corresponding block partition $B_{i,j}$ of the lowest $3L' + 1$ subbands of a component of texture.

The process of embedding of a given DEM coefficient in the Y or Cr/Cb component of the transformed texture is elaborated by the flowchart given in Fig. 3. Since one coefficient is embedded per $t/2^{(L-L')} \times t/2^{(L-L')}$ block, each $B_{i,j}$ is repartitioned into as many sub-blocks (b_k) as are the number of bits used to represent a single transformed DEM coefficient. Each of b_k from $B_{i,j}$ would carry the k^{th} bit β_k of the transformed DEM coefficient $d_{i,j}$. Embedding depends on a key generating a pseudo-random matrix W with entries from the set $\{\alpha, -\alpha\}$. The matrix W has the same size as of b_k , i.e. $W = [y_{uv}]_{n \times n}$, where $y_{uv} \in \{\alpha, -\alpha\}$. The scalar, α , is referred to as the strength of embedding. If the bit to embed, β_k , is '1', then matrix W is added to the matrix b_k , otherwise W is subtracted from b_k . The result is a new matrix b'_k which replaces b_k as a sub-block in the embedded block, $B'_{i,j}$, of the marked carrier C' . Two factors are important over here. First is the DWT level (L') of DEM before embedding which is a trade off between the final texture quality and its DEM quality. At the decoding end the quality of the DEM would depend on the difference between L and L' . The larger the difference, $(L - L')$, higher will be the quality and vice versa. Second is the value of α since a larger α means high degradation of the embedded texture. This second factor is, however of secondary importance since the embedded message (M) after recovery, will be used to correct any loss in texture quality. So, no matter how much degradation is there, the reconstructed texture would be of high quality.

3.4 The decoding step

The above coded image can be utilized like any other JPEG2000 image and sent across any communication channel. The decoding is more or less converse to the above process. Just before the inverse DWT stage of the JPEG2000 decoder the DEM can be extracted from C' , using the above mentioned partitioning scheme, i.e. $B'_{i,j}$ blocks and their b'_k sub-blocks. Fig. 4 shows the flowchart for the recovery of a DEM coefficient $d_{i,j}$ from sub-blocks b'_k of $B'_{i,j}$ and eventual reconstruction of $b_{i,j}$ and ultimately $B_{i,j}$. A given k^{th} sub-block b'_k and

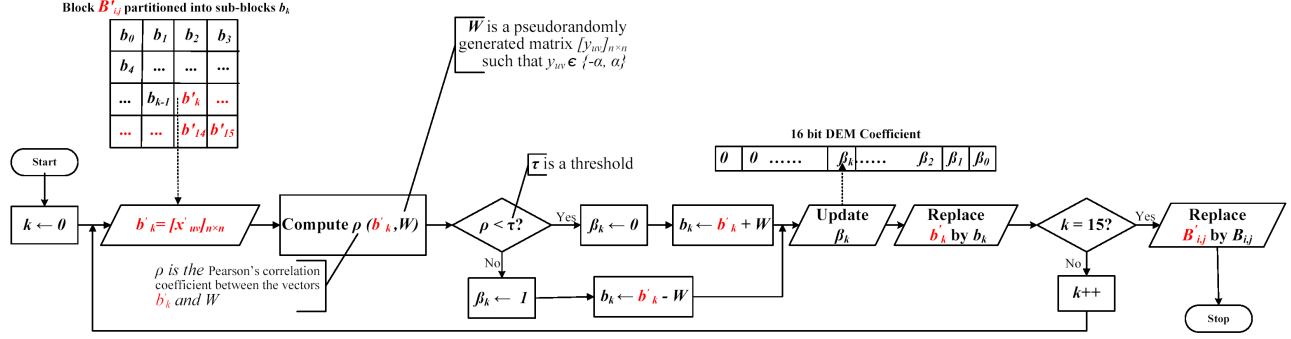


Figure 4: Flowchart showing the recovery process of the DEM coefficient $d_{i,j}$ and the corresponding block $B_{i,j}$ from the SS embedded texture block partition $B'_{i,j}$ of the lowest $3L' + 1$ subbands of an embedded texture component.

also the matrix W can be treated as a column/row vector and the Pearson's correlation coefficient, ρ , can be computed thereof. If ρ is less than certain threshold, τ then the embedded bit β was a '0', otherwise it was a '1'. Once this is determined then it is obvious that what were the entries of b_k , i.e. if β is '0' then add W to b'_k , otherwise subtract W from b'_k .

3.5 The reconstruction step

Now comes the reconstruction phase wherein by the application of 0-padding one can have $L+1$ and $L'+1$ different approximation images of texture and DEM, respectively. And this is where one achieve the scalability goal. Our method is based on the assumption that it is not necessary that all the data is available for reconstruction, i.e. only a part of C' is on hand. This is one of the main advantages of the method since the DEM and texture can be reconstructed with even a small subset of the coefficients of the carrier. The resolution scalability of wavelets and the synchronized character of our method enable a 3D visualization even with fewer than original resolution layers as a result of partial or delayed data transfer. The method thus enables to effect visualization from a fraction of data in the form of the lowest subband, of a particular resolution level since it is always possible to stuff 0's for the higher bands. The idea is to have a 3D visualization utilizing lower frequency subbands at level L'' , say, where $L'' \leq L$. For the rest of $L - L''$ parts one can always pad a 0 for each of their coefficient. The inverse DWT of the 0-stuffed transform components will yield what is known as image of approximation of level L'' . Before this, data related to DEM must be extracted whose size depends on both L'' and L' . Thus if $L' \leq L''$ one will always have the entire set of the embedded DEM coefficients since all of them will be extractable. We would have a level 0 approximate final DEM after DWT^{-1} , of the highest possible quality. On the other hand if $L' > L''$, one would have to pad 0's for all coefficients of higher $L' - L''$ subbands of transformed DEM before DWT^{-1} that would result in a level L'' -approximate DEM of an inferior quality.

4. RESULTS AND ANALYSIS

To explain the effectiveness of our method, we chose the example texture/DEM pair given in Fig. 5. The DEM is shown (Fig. 5.a) in the form of a 32×32 gray scale image wherein the whiteness determine the height of the altitude. The corresponding texture (Fig. 5.b) has a size of 3072×3072 pixels implying one DEM coefficient per 96×96 texture block, i.e. $t = 96$. For the purpose of comparison a 256×256 pixels portion, at (1000, 1500) coordinates, is magnified as shown in Fig. 5.c. Fig. 5.d illustrates a 3D view obtained with the help of the texture/DEM pair. We chose to subject the texture to reversible JPEG2000 encoding at $L = 4$ that would give us five possible resolutions (13 subbands) based on 1, 4, 7, 10 lowest frequency or all of the 13 sub

For fully synchronous embedding all the 13 subbands of the selected component were utilized in embedding and thus the DEM was subjected to lossless DWT at level $L' = L = 4$ to give 13 subbands. The embedding block, $B_{i,j}$, had then a size of 96×96 for the dissimulation of one 16 bit transformed DEM coefficient which means a 24×24 subblock (b_k) per DEM bit. Since b_k is large enough, the needed strength/amplitude (α) of

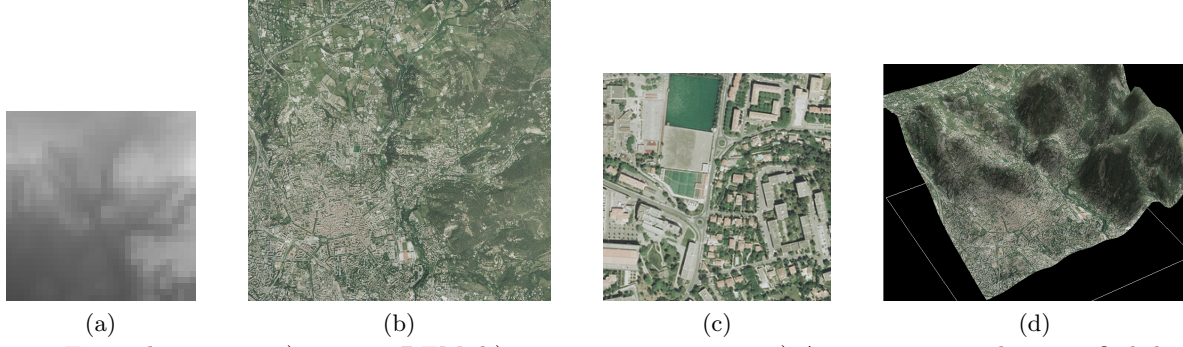


Figure 5: Example images: a) 32×32 DEM, b) 3072×3072 Texture, c) A 256×256 pixels magnified detail at (1000, 1500) coordinates, d) A corresponding 3D view.

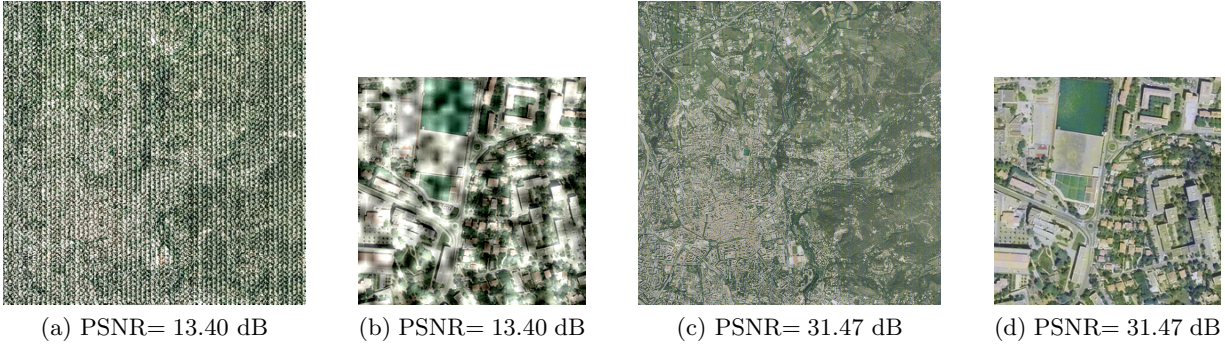


Figure 6: Coded texture after embedding in a) Y-plane ($\alpha = 64$), b) its 256×256 pixels magnified detail, c) Cr plane ($\alpha = 16$) and d) its 256×256 pixels magnified detail.

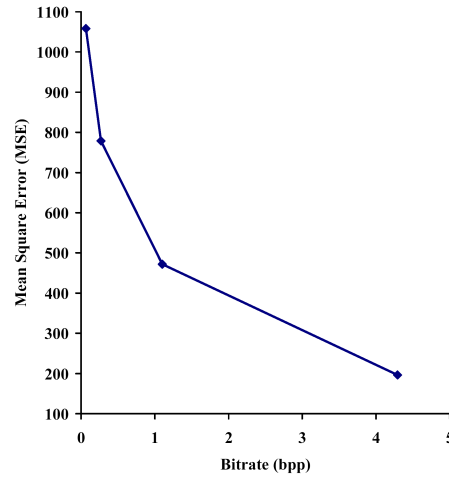


Figure 7: Variation in texture quality as a function of its bitrate for a DEM error of 3.18 m .

embedding is smaller and the quality of luminance/chrominance component will not be deteriorated much. For our example it was observed that 100% successful recovery of the embedded message is realized when $\alpha = 2$, for embedding chrominance component (Cr/Cb), as against $\alpha = 8$ for embedding luminance component(Y). The overall quality of the embedded texture was observed to be better (44.39 dB) for Cr/Cb than for Y (26.53 dB). This quality difference is not that important, however, since the original texture is almost fully recoverable from the embedded texture. On the other hand there is a risk that embedding in Cr/Cb may eventually inflate the size of the coded image. Upon the reintroduction of the marked Y or Cr/Cb to the JPEG2000 pipeline we get our embedded texture image in JPEG2000 format. From this image one can have 5 different approximation

images for both the DEM and the texture. A level- l ($l \leq L$) approximate image is the one that is constructed with $(1/4^l) \times 100$ percent of the total coefficients that corresponds to the available lower $3l + 1$ subbands. For example, level-0 approximate image is constructed from all the coefficients and level-2 approximate image is constructed from 6.12% of the count of the initial coefficients. Since the embedded data is removable, one gets the highest possible qualities for all the texture approximations but not for the DEM, as its quality depends on our embedding strategy.

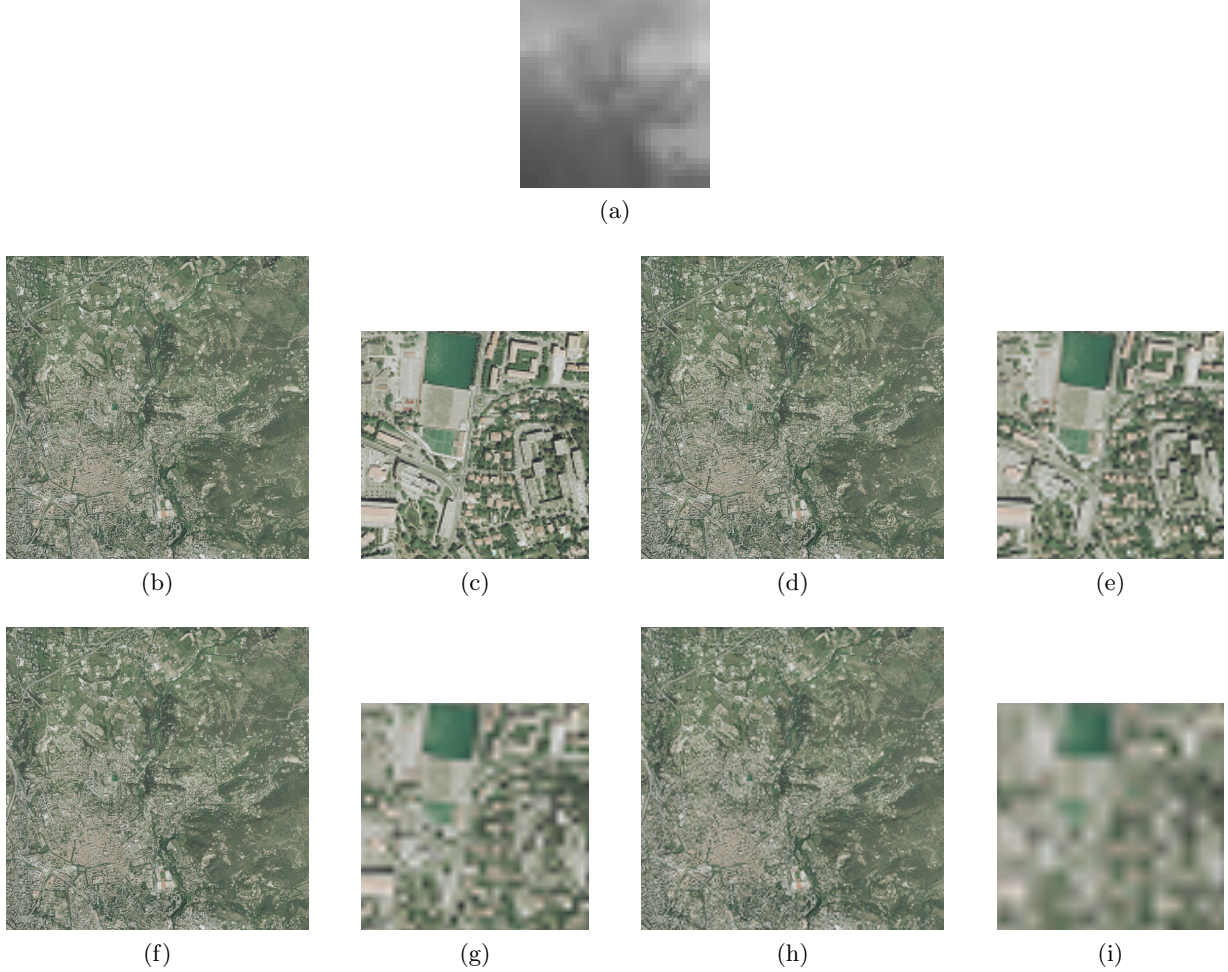


Figure 8: Approximation images corresponding to 3.18 m DEM error when level- L' transformed DEM is embedded in lowest $3L' + 1$ subbands of level-4 transformed Cr texture: a) Level-1 approximate DEM having 3.18 m error, b) Level-1 approximate texture (4.29 bpp) with $L' = 4$, c) its 256×256 pixels magnified detail, d) Level-2 approximate texture (1.1 bpp) with $L' = 3$, e) its 256×256 pixels magnified detail, f) Level-3 approximate texture (0.27 bpp) with $L' = 2$, g) its 256×256 pixels magnified detail, h) Level-4 approximate texture (0.066 bpp) with $L' = 1$ and i) its 256×256 pixels magnified detail.

The sensitive nature of DEM compels us to avoid too much loss in its quality. For improved DEM quality one had to adjust the synchronization and rather than persisting with $L' = 4$ we went for $L' < 4$, i.e. which meant exclusion of $3(4 - L')$ highest frequency subbands, of the carrier, from embedding; the synchronization was now maintainable between all the $3L' + 1$ subbands of DEM and the subset $3L' + 1$ lowest subbands of the carrier. Obviously dimensions of $B_{i,j}$ and b_k were dyadically reduced by a factor of $2^{4-L'}$ which led to an increase in the value of α and an eventual degradation in the quality of the coded texture. As described in Section 3.4, the most important step of our approach is to adapt the synchronization. In other words, the objective of the step is to find the lower bound for L' . This bound depends on the texture to DEM size ratio and also on the fact that how

much distortion in the carrier is reversible, i.e. bounds of α . For our example, the lowest quality for the coded texture - 13.40 dB with $\alpha = 64$ for Y carrier (Fig. 6.a-b) and 31.47 dB with $\alpha = 16$ for Cr carrier (Fig. 6.c-d) - was observed at $L' = 1$. The reason is that all the information had to be embedded in just 4 lowest energy and smallest subbands implying a size of 3×3 for b_k with $\alpha = 64$ for Y carrier and $\alpha = 16$ for Cr/Cb carrier. Beyond this bound ($L' = 0$, i.e. spatial domain) the block size did not allow for the recovery of the embedded message.

To judge the DEM quality, root mean square error (RMSE) in meters (m), as explained in Section 3.2, was adopted as a measure. With our approach, by using the example given in Fig. 5, the RMSE of the five possible DEM approximations, i.e. level 0, 1, 2, 3 and 4 were found to be 0 m , 3.18 m , 6.39 m , 11.5 m and 22.88 m , respectively. For full synchronization, as presented in Ref. 4, the worst DEM quality, 22.88 m , is resulted when one go for a 3D visualization from level-4 approximate texture as the corresponding DEM will also be 4-approximate. But even with one step adjustment this quality is twice improved and for $L' = 3$, both the 3- and 4- approximate texture images have 3- approximate DEM with RMSE = 11.5 m . Go a step further and maximum DEM error will be reduced to 6.39 m . Hence with adaptive synchronization one can have a high quality DEM even for very low quality texture or, more precisely, the same quality DEM for all the approximation images at levels $\geq L'$. Fig. 7 shows the trend in texture quality for a given DEM error (3.18 m) as a function of bitrate. Images corresponding to a DEM error of 3.18 m are shown in Fig. 8 and the resultant visualization is illustrated in Fig. 9. It can be seen that with even a tiny fraction of the total coefficients (as low as 0.40% ,i.e. Fig. 8h-i and Fig. 9.e), a fairly commendable visualization can be realized.

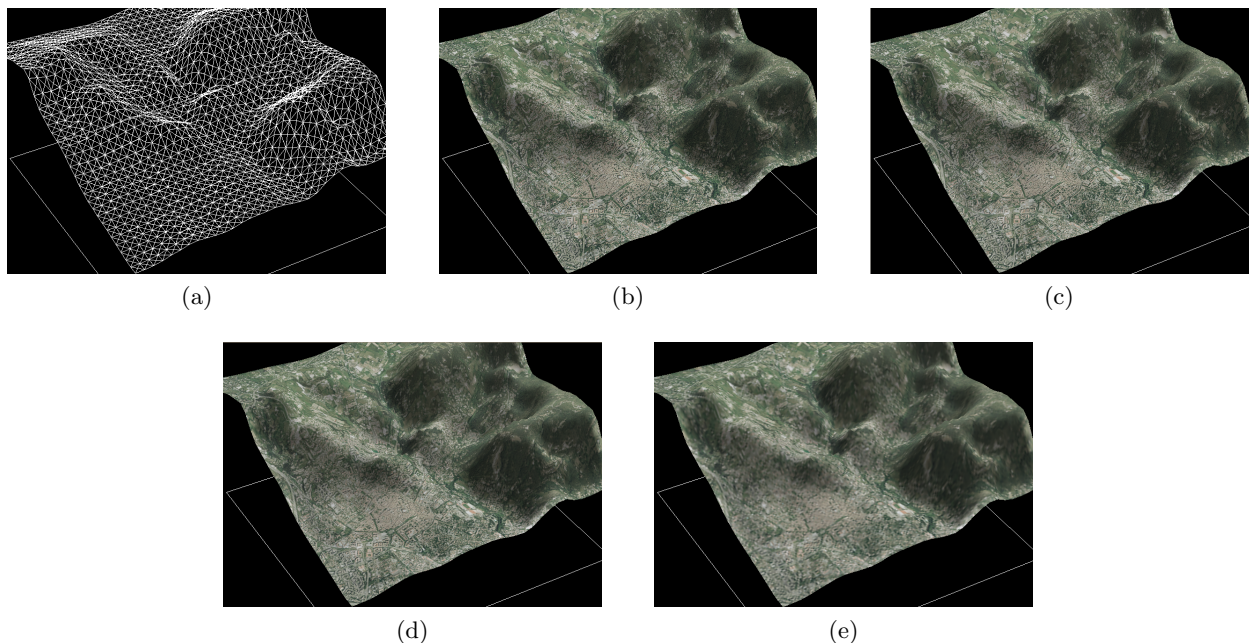


Figure 9: 3D visualization corresponding to Fig. 8: a) Level-1 approximate DEM having 3.18 m error, b) Level-1 approximate texture (4.29 bpp) with $L' = 4$, c) Level-2 approximate texture (1.1 bpp) with $L' = 3$, d) Level-3 approximate texture (0.27 bpp) with $L' = 2$, e) Level-4 approximate texture (0.066 bpp) with $L' = 1$.

The maximum error in RMSE of the DEM tolerable by an observer is a function of the distance of the observer, i.e. the viewpoint D . This threshold is mainly dependent on the human visual system (HVS) and for this reason, the analysis given in Section 3.2 is extremely useful to adapt the synchronization.

5. CONCLUSION

By taking advantage of the multiresolution scalability of JPEG2000, we unified the disparate 3D terrain data synchronously by employing an adaptive SS scalable data hiding strategy. The reversibility of the method enabled us to improve the highest possible quality of the texture while the adaptive synchronization ameliorated

the DEM quality for a particular texture quality. It can be argued that the adapting depends on the initial block size and with the development of technology, adaptation may no longer be possible. But the technology is not improving for the DEM only as texture images are also becoming more and more refined with the passage of time. Hence the block size may vary but there is every likelihood that the texture to DEM size ratio may still remain significantly large to allow for the adaptive synchronization.

In the continuation of our work, it will be prudent to utilize the robustness of the SS embedding to our advantage. In the near future, one would like to realize a 3D visualization if a small patch is cropped from the encoded image. The issue would then be the extent of extractability of the corresponding DEM data from the coded texture patch. Another issue that needs immediate attention, is about the focus of the viewpoint. One need highest resolution at the focused area but not necessarily outside if the memory and computer resources are not sufficient. This problem is specially noticeable in case of Geo-browsers like Google Earth or NASA's World Wind. A strategy can be to use fully reconstructed texture at focus and tile it with approximation images at the periphery. This would however require smoothening of the tile boundaries.

ACKNOWLEDGMENTS

This work is in part supported by the Higher Education Commission (HEC) of Pakistan and in part by the French National Project VOODOO of the ANR Content and Interaction.

REFERENCES

- [1] Losasso, F. and Hoppe, H., "Geometry Clipmaps: Terrain Rendering Using Nested Regular Grids," *ACM Trans. Graph.* **23**(3), 769–776 (2004).
- [2] Pajarola, R. and Gobbetti, E., "Survey on Semi-Regular Multiresolution Models for Interactive Terrain Rendering," *The Visual Computer* **23**, 583–605 (June 2007).
- [3] ISO/IEC, [ISO/IEC 15444-1: Information technology, JPEG2000 image coding system, Part 1: Core coding system], ISO Central Secretariat: CH-1211 Geneva, Switzerland (2004).
- [4] Hayat, K., Puech, W., and Gesquière, G., "Scalable 3D Visualization Through Reversible JPEG2000-Based Blind Data Hiding," *IEEE Trans. on Multimedia* **10**(7) (2008).
- [5] Hayat, K., Puech, W., and Gesquière, G., "A Lossy JPEG2000-Based Data Hiding Method for Scalable 3D Terrain Visualization," in [Proc. 16th European Signal Processing Conference (EUSIPCO 08)], EURASIP, Lausanne, Switzerland (August 2008).
- [6] Hayat, K., Puech, W., and Gesquière, G., "Scalable Data Hiding for Online Textured 3D Terrain Visualization," in [Proc. ICME'08, IEEE International Conference on Multimedia & Expo], 217–220 (June 2008).
- [7] Cox, I. J., Miller, M. L., and Bloom, J. A., [Digital Watermarking], Morgan Kaufmann Publishers (2002).
- [8] Meerwald, P. and Uhl, A., "A Survey of Wavelet-Domain Watermarking Algorithms," in [Proc. SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents III], **4314**, 505–516, SPIE, IS&T, San Jose, CA, USA (January 2001).
- [9] Kundur, D., "Improved Digital Watermarking Through Diversity and Attack Characterization," in [Proc. ACM Workshop on Multimedia Security'99], 53–58 (Oct. 1999).
- [10] Kundur, D. and Hatzinakos, D., "Digital Watermarking Using Multiresolution Wavelet Decomposition," in [Proc. IEEE International Conference on Acoustic, Speech and Signal Processing (IEEE ICASSP 98)], **5**, 2969–2972 (May 1998).
- [11] Xia, X. G., Boncelet, C. G., and Arce, G. R., "A Multiresolution Watermark for Digital Images," in [Proc. IEEE International Conference on Image Processing (IEEE ICIP 97)], 548–551 (Oct. 1997).
- [12] Kundur, D. and Hatzinakos, D., "A Robust Digital Image Watermarking Scheme Using the Wavelet-Based Fusion," in [Proc. IEEE International Conference on Image Processing (IEEE ICIP 97)], **1**, 544–547 (Oct. 1997).
- [13] Su, P. C., Wang, H. J., and Kuo, C. C. J., "An Integrated Approach to Image Watermarking and JPEG-2000 Compression," *Journal of VLSI Signal Processing Systems, Special Issue on Multimedia Signal Processing* **27**, 35–53 (Jun. 1997).

- [14] Meerwald, P., "Quantization watermarking in the JPEG2000 coding pipeline," in [*Proc. Communications and Multimedia Security Issues of The New Century, IFIP TC6/TC11 Fifth Joint Working Conference on Communications and Multimedia Security, CMS '01*], Steinmetz, R., Dittmann, J., and Steinebach, M., eds., 69–79, Kluwer Academic Publishing (May 2001).
- [15] Piva, A., Bartolini, F., and Caldelli, R., "Self Recovery Authentication of Images in The DWT Domain," *Int. J. Image Graphics* **5**(1), 149–166 (2005).
- [16] Liu, J. L., Lou, D. C., Chang, M. C., and Tso, H. K., "A Robust Watermarking Scheme Using Self-Reference Image," *Computer Standards & Interfaces* **28**, 356–367 (2006).
- [17] Kong, X., Liu, Y., Liu, H., and Yang, D., "Object Watermarks for Digital Images and Video," *Image and Vision Computing* **22**, 583–595 (2004).
- [18] Noore, A., Singh, R., Vatsa, M., and Houck, M. M., "Enhancing Security of Fingerprints through Contextual Biometric Watermarking," *Forensic Science International* **169**(2–3), 188–194 (2007).
- [19] Xiang, S. and Kim, H.-J., "Geometrically Invariant Image Watermarking in the DWT Domain," in [*WISA*], Kim, S., Yung, M., and Lee, H.-W., eds., *Lecture Notes in Computer Science* **4867**, 76–90, Springer (2007).
- [20] Lin, W.-H., Horng, S.-J., Kao, T.-W., Fan, P., Lee, C.-L., and Pan, Y., "An Efficient Watermarking Method Based on Significant Difference of Wavelet Coefficient Quantization," *IEEE Trans. on Multimedia* **10**(5), 746–757 (2008).
- [21] Schlauweg, M., Pröfrock, D., and Müller, E., "JPEG2000-Based Secure Image Authentication," in [*MM&Sec '06: Proceedings of the 8th Workshop on Multimedia and Security*], 62–67, ACM, New York, NY, USA (2006).
- [22] Agreste, S., Andaloro, G., Prestipino, D., and Puccio, L., "An Image Adaptive, Wavelet-Based Watermarking of Digital Images," *Journal of Computational and Applied Mathematics* **210**(1–2), 13–21 (2007).
- [23] Dawei, Z., Guanrong, C., and Wenbo, L., "A Chaos-Based Robust Wavelet-Domain Watermarking Algorithm," *Chaos, Solitons & Fractals* **22**(1), 47–54 (2004).
- [24] Uccheddu, F., Corsini, M., and Barni, M., "Wavelet-Based Blind Watermarking of 3D Models," in [*MM&Sec '04: Proceedings of the 2004 workshop on Multimedia and security*], 143–154, ACM, New York, NY, USA (2004).
- [25] Yin, K., Pan, Z., Shi, J., and Zhang, D., "Robust Mesh Watermarking Based on Multiresolution Processing," *Computers & Graphics* **25**(3), 409–420 (2001).
- [26] Kim, J. and Ra, J., "A Real-Time Terrain Visualization Algorithm Using Wavelet-Based Compression," *The Visual Computer* **20**(2–3), 67–85 (2004).
- [27] Gerlek, M. P., [*The "GeoTIFF Box" Specification for JPEG 2000 Metadata - DRAFT version 0.0*], LizardTech, Inc. 1008 Western Ave Suite 200 Seattle, WA 98104 USA (April 2004).
- [28] Lake, R., Burggraf, D., Kyle, M., and Forde, S., [*GML in JPEG 2000 for Geographic Imagery (GMLJP2) Implementation Specification*], no. OGC 05-047r2, Open Geospatial Consortium (OGC) (2005).
- [29] Vepakomma, U., St-Onge, B., and Kneeshaw, D., "Spatially Explicit Characterization of Boreal Forest Gap Dynamics Using Multi-Temporal Lidar Data," *Remote Sensing of Environment* **112**(5), 2326–2340 (2008).
- [30] Chauve, A., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M., and Puech, W., "Processing Full-Waveform Lidar Data: Modelling Raw Signals," in [*Proc. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*], **36**(Part 3/W52), 102–107 (2007).
- [31] Smith, G. and Atchison, D., [*The Eye and Visual Optical Instruments*], Cambridge University Press, New York, NY, USA (1997).
- [32] Luebke, D., Reddy, M., Cohen, J. D., Varshney, A., Watson, B., and Huebner, R., [*Level of Detail for 3D Graphics*], Morgan Kaufmann Publishers, San Francisco, CA, USA (July 2002).
- [33] Puech, W., Bors, A. G., Pitas, I., and Chassery, J.-M., "Projection Distortion Analysis for Flattened Image Mosaicing from Straight Uniform Generalized Cylinders," *Pattern Recognition* **34**(8), 1657–1670 (2001).